Problems of Information Retrieval in Polyphonic Music

Donald Byrd

CIIR, University of Massachusetts

9 June 2000

Joint work with Tim Crawford, Jeremy Pickens, and Matthew Dovey

Overview

1. Background and Review

   Why is Music IR Useful? Why is It Hard?

2. Music Perception and Music IR

3. Problems Matching Musical Data

4. Conclusions

Background

- First published work on music IR goes back to mid-1960's
- Now, interest in music IR is exploding: almost every relevant recent conference has had papers on music retrieval and/or digital music libraries
- First major grant for music-IR research, to us, was recently funded

Background

- Music IR is still a very immature field
  - No survey of user needs has ever been done
  - No one we know of has relevance judgements for any music collection and set of queries
  - Indexing is as vital for music as it is for text—but requirements quite different, and first published research on indexing music dates back less than five years
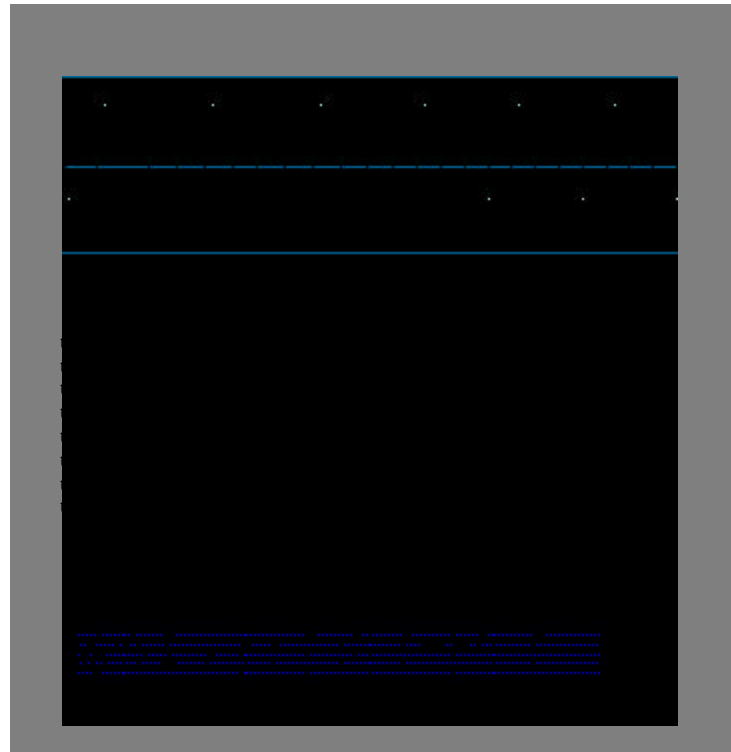- Music IR now decades behind text IR

Audio (e.g., CD, MP3)

Time-stamped Events
(e.g., MIDI file)

Music Notation

	Audio	*Time-stamped Events*	*Music Notation*
Common examples	CD, MP3 file	Standard MIDI File	Sheet music
Unit	Sample	Event	Note, clef, lyric, etc.
Explicit structure	none	little (partial voicing information)	much (complete voicing information)
Avg. rel. storage	2000	1	10
Convert to left	-	easy	OK job: easy
Convert to right	1 note: pretty easy other: hard? Very hard?	OK job: fairly hard	-
Ideal for	music bird/animal sounds sound effects speech	music	music

Review: Why is Music IR Research Worthwhile?

- It's Useful
  - Continuum from music lovers to music researchers
  - Intellectual-Property Rights people
- It's Hard
  - unique challenges vs. text or image IR

Review: Why is Music IR Useful?

- Music Information Needs

    1. "Name That Tune": identify a specific piece (known-item search)

    2. Find Versions: find recordings/music of versions (a.k.a. arrangements, covers, transcriptions) of a specific piece

    3. Find by Style: find pieces in same style as query

    4. Find Similar: find pieces similar to query

- Needs of music lovers to music researchers

- Needs of Intellectual-Property Rights people

    – Especially #4 via routing, not conventional IR!

Review: Why is Music IR Hard?

1. Units of meaning: not clear anything in music is analogous to words *(all representations)*

2. Polyphony: parallel independent voices, something like characters in a play *(all representations)*

3. Indexing: both of above make indexing very difficult *(all representations)*

4. Recognizing notes *(audio only)*

5. Other reasons
   - Musician-friendly I/O is difficult
   - Diversity: of styles of music, of people interested in music

MARLENE. What I fancy is a rare steak. Gret?

ISABELLA. I am of course a member of the / Church of England.*

GRET. Potatoes.

MARLENE. *I haven't been to church for years. / I like Christmas carols.

ISABELLA. Good works matter more than church attendance.

--Caryl Churchill: "Top Girls" (1982), Act 1, Scene 1

Performance (time goes from left to right):

M: What I fancy is a rare steak. Gret?   I haven't been...

I: I am of course a member of the Church of England.

G: Potatoes.

Review: Units of Meaning (Problem 1)

- Not clear anything in music is analogous to words
  - No explicit delimiters (like Chinese)
  - Experts don't agree on "word" boundaries (unlike Chinese)
- Are *notes* like words?
- No. Relative, not absolute, pitch is important
- Are *pitch intervals* like words?
- No. They're too low level: more like characters
- Are *pitch-interval sequences* like words?
- In some ways, but
  - Ignores note durations
  - Ignores relationships between voices (harmony)
  - Probably little correlation with semantics

Review: Parallel Voices in Music (Problem 2)

J.S. Bach: "St. Anne" Fugue, beginning

The Four Parameters of Notes

- Four basic parameters of a definite-pitched musical note
  - *pitch:* how high or low the sound is: perceptual analog of frequency
  - *duration:* how long the note lasts
  - *loudness:* perceptual analog of amplitude
  - *timbre* or tone quality

Music Perception and Music IR

- Salience is affected by texture, loudness, etc.
  - Example(?)
- Streaming effects and cross-voice matching
  - produced by timbre: Wessel's illusion
  - produced by register: Telemann example
- Octave identities, timbre and texture
  - Beethoven "Hammerklavier" Sonata example
  - Affects pitch-interval matching

Music IR and the Four Parameters of Notes

- What is the relative weight of information carried by each of four parameters in a given style of music?

- For mainstream Western music, average maybe pitch 50%, rhythm 40%, timbre and dynamics 10%

- Pitch occurs in both the horizontal (melodic) and vertical (harmonic) dimensions

- Rhythm is not just strings of durations: it involves meter and accent patterns as well.

  – Caveat: factors are not independent.

Music IR and the Four Parameters of Notes

- Much of information in music is not in pitch, certainly not only in horizontal (melodic) pitch.

- Yet almost all music-IR work to date has done only melodic pitch matching—and with a fair amount of success.

- But almost all music-IR work has also had severe restrictions.

Music IR and the Four Parameters of Notes

- Database size
  - Work so far: moderate-sized databases (c. 10K small documents)
  - Music holdings of the Library of Congress: over 10M items, including over 6M pieces of sheet music and tens of thousands of scores of operas, symphonies, etc.

- Complexity of music
  - Work so far: relatively simple music, usually monophonic
  - A Mozart symphony might use 12 voices; Stravinsky's *Le Sacre du Printemps* uses up to c. 38; film/TV music is similar

- Will melodic pitch alone be adequate for large databases and complex music?

Wanted Beethoven, Found Dvorak

- Barlow and Morgenstern's *Dictionary of Musical Themes* has 10,000 themes and a 100-page pitch-interval index
- I looked in index for the main theme of the "Ode to Joy" of Beethoven's 9th Symphony
- It has an entry for the 1st six notes of the theme, but points only to a Dvorak piece that sounds completely different!
- Main cause of the false positive: the index ignores rhythm
- Cause of the false negative: most instances of the theme in the Beethoven, especially more salient ones, involve trivial melodic ornamentation: subdivision of the first note into repeated notes (the version I searched for)

Melodic Confounds

- Repeated notes, rests, and grace notes "can confound both contour and intervallic-profile comparisons."
- In Barlow & Morgenstern's *Dictionary of Musical Themes*
  - Repeated notes: 40-50% of themes
  - Rests: 35%
  - Grace notes: 15%
- In the *Real Vocal Book*
  - Repeated notes: 53-65%
  - Rests: 67%
  - Grace notes: none

Problems Matching Musical Data

- Objective: inherent in complexity of music
  - - Replacement (note-for-note): e.g., tonal answer, mutation
  - - Melodic ornamentation (simple/subdivision or complex/insertion)
  - - Melodic simplification (deletion)
  - - Melody crossing voices (distributed matching)
- Subjective: errors of cognition and memory
  - - Version remembered is simplified
  - - Version remembered mixes voices
  - - Version remembered is incorrect in some other way
- Other: errors in input
  - - Outright mistakes in query and/or database
  - - Performance-related issues: affected by style and genre

20#

Prospects for Solutions
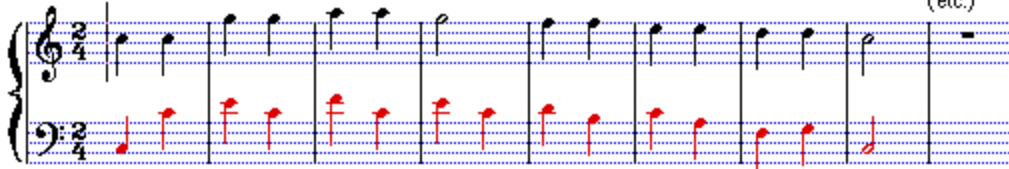
- These problems are all common now

- How good are prospects for solving them?

  - Objective problems are inherent in music => will remain common

  - Subjective problems and mistakes by user result from human nature => will remain common

  - Outright mistakes from conversion are common now in OMR, and much more in AMR, systems

  - May become less common in OMR, but "AMR is orders of magnitude more difficult than OMR" => must assume they'll remain common in AMR, at least for many years

Nightingale and Extra Notes (Problem 2)
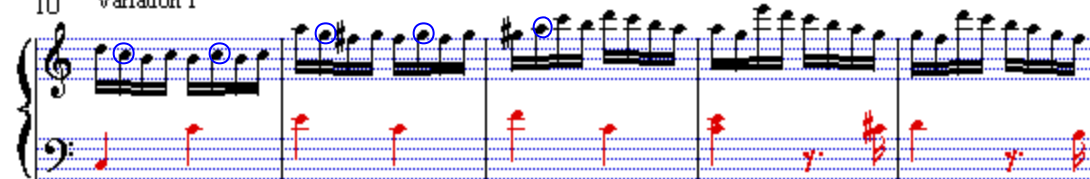
Mozart: Variations on "Ah, vous dirai-je, Maman" for piano, K. 265, Theme & Var. 1

Nightingale and Parallel Voices (Problem 2)

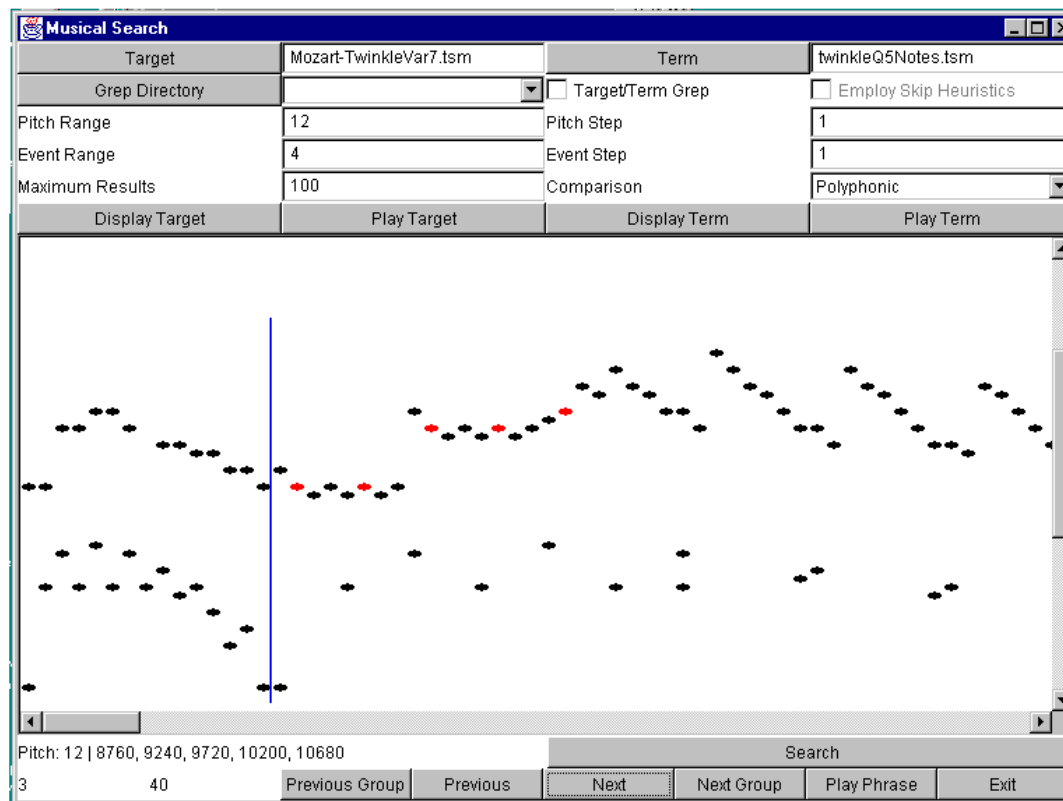Mozart: Variations on "Ah, vous dirai-je,
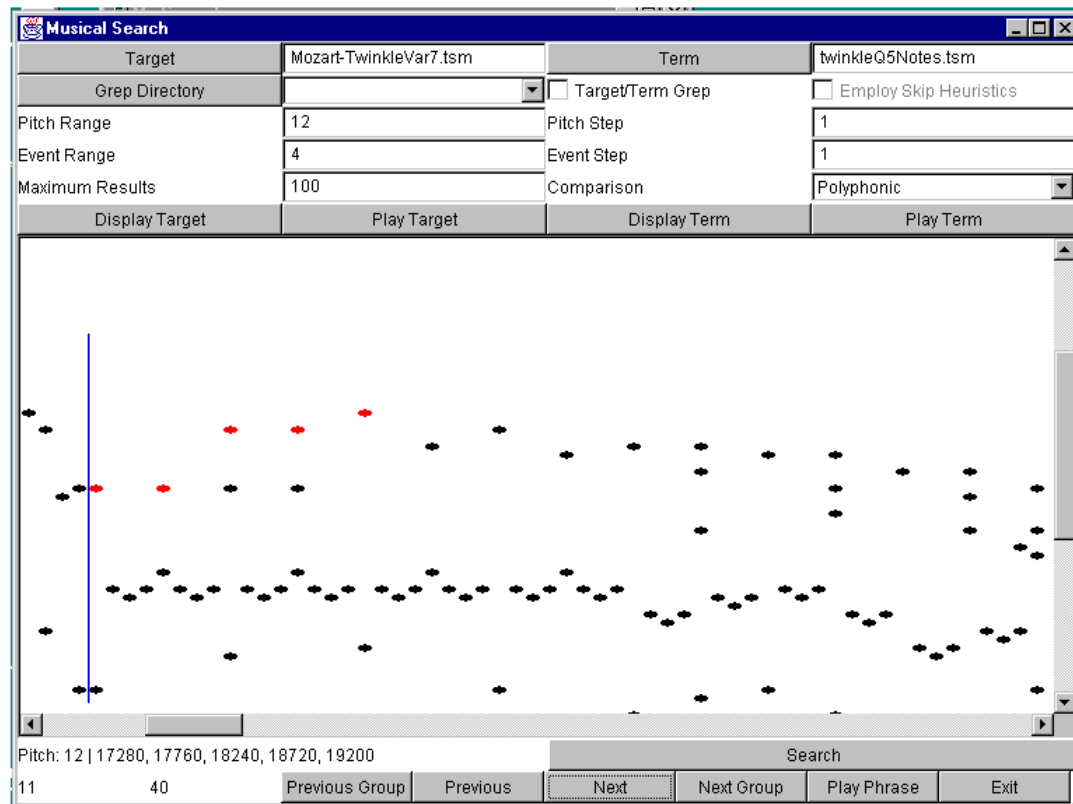Maman" for piano, K. 265, Variation 2

Mozart: Variations on "Ah, vous dirai-je,
Maman" for piano, K. 265, Theme & Var. 1

2-D Pattern Matching in JMS and Parallel Voices

Mozart: Variations on "Ah, vous dirai-je,
Maman" for piano, K. 265, Variation 2

Conclusions

- Music IR is decades behind text IR, but what did you expect?

- Seem to be making progress on difficult problems
  - Funding agencies and corporations are finally spending $$
  - Research community is working on indexing, polyphony, etc.

- But how do we know any of this is any good?
  - Community is thinking about testbeds (databases, relevance judgements) and Cranfield-model evaluation (like TREC), etc., so we'll "know"
  - Music IR 2000!

- Don't need perfection to be useful: cf. text IR!